INVESTOR IN PEOPLE

#2

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ



# PRIORITY DOCUMENT

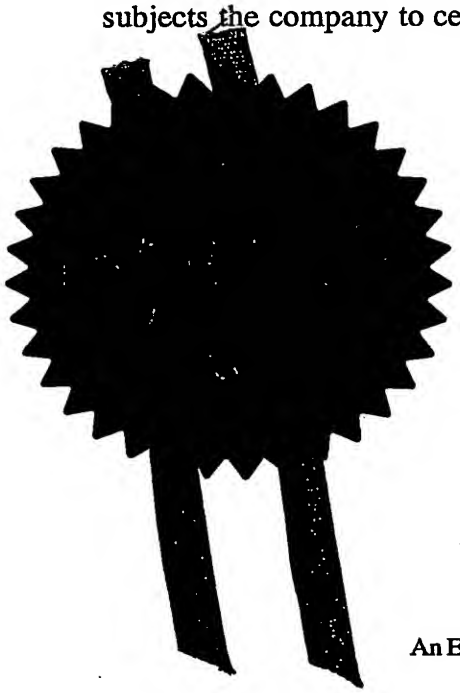
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.



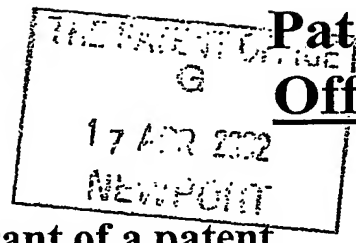
Signed P. Mahoney

Dated 29 April 2003

Patents Form 1/77

*Patents Act 1977*
*(Rule 16)*

# The Patent Office




1/77

17APR02 E711736-1 C50092
P01/7700 0.00-0208768.2

## Request for grant of a patent

*(see the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)*

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

| | | |
|---|---|---|
| 1. | Your reference | PCB/P089103PGB |
| 2. | Patent application number *(The Patent Office will fill in this part)* | 0208768.2 — 17 APR 2002 |
| 3. | Full name, address and postcode of the or of each applicant *(underline all surnames)* | THE UNIVERSITY OF NEWCASTLE<br>6 KENSINGTON TERRACE<br>JESMOND<br>NEWCASTLE UPON TYNE<br>NE1 7RU |
| | Patents ADP number *(if you know it)* | 617916 2001 |
| | If the applicant is a corporate body, give the country/state of its incorporation | UNITED KINGDOM |
| 4. | Title of the invention | DNA POLYMERASES |
| 5. | Name of your agent *(if you have one)* | Marks & Clerk |
| | "Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)* | 43 Park Place<br>Leeds<br>LS1 2RY |
| | Patents ADP number *(if you know it)* | 18013 ✓ |

| | | | | |
|---|---|---|---|---|
| 6. | If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number | Country | Priority application number *(if you know it)* | Date of filing *(day/month/year)* |
| 7. | If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application | Number of earlier application | | Date of filing *(day/month/year)* |
| 8. | Is a statement of Inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*<br>*a) any applicant named in part 3 is not an inventor, or*<br>*b) there is an inventor who is not named as an applicant, or*<br>*c) any named applicant is a corporate body.*<br>*See note (d))* | YES | | |

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

| | |
|---|---|
| Continuation sheets of this form | - |
| Description | 23 |
| Claim(s) | 3 |
| Abstract | - |
| Drawing(s) | 4 + 4 |

10. If you are also filing any of the following, state how many against each item.

| | |
|---|---|
| Priority documents | - |
| Translations of priority documents | - |
| Statement of Inventorship and right to grant of a patent *(Patents Form 7/77)* | - |
| Request for preliminary examination and search *(Patents Form 9/77)* | 1 |
| Request for substantive examination *(Patents Form 10/77)* | - |
| Any other documents *(Please specify)* | - |

11. I/We request the grant of a patent on the basis of this application.

Signature MARKS & CLERK Date 16/04/02

12. Name and daytime telephone number of person to contact in the United Kingdom MR P. C. BANFORD – 0113 3895606

**Warning**

*After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.*

**Notes**

*a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.*
*b) Write your answers in capital letters using black ink or you may type them.*
*c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*
*d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.*
*e) Once you have filled in the form you must remember to sign and date it.*
*f) For details of the fee and ways to pay please contact the Patent Office.*

# DNA POLYMERASES

The present invention relates to archaeal DNA polymerase variants and their use in the amplification of DNA.

Polymerase chain reaction (PCR) is a method whereby a sequence of DNA may be selectively amplified to produce a large sample that may be readily analyzed. A solution containing the DNA to be amplified, together with free bases, a polymerase enzyme and primers that bind to opposite ends of the two strands of the DNA segment to be replicated, is heated to break the bonds between the strands of DNA. When the solution cools, the primers bind to the separated strands and the polymerase builds a new strand by joining free bases to the primers thereby producing a new strand that is restricted solely to the desired segment. PCR enables billions of copies of a small piece of DNA to be produced in several hours.

Heat stable polymerases are required for this process and one of the most commonly used is Taq DNA polymerase from *Thermus aquaticus*. However, this enzyme does not possess a 3'-5' exonuclease III function, commonly referred to as "proofreading activity". This function removes bases that are mismatched at the 3' end of a primer-template duplex. The inability of Taq DNA polymerase to carry out this function results in it being prone to base incorporation errors.

Archaeal DNA polymerases are thermally stable and demonstrate proofreading activity. However, native archaeal DNA polymerases are inhibited by deoxyuracil. Archaeal DNA polymerases have a "read-ahead" function specifically for uracil. This template

checking activity scans the template ahead of the replication fork for the presence of uracil and stalls polymerisation when uracil is encountered. Thus, the presence of deoxyuracil in DNA causes amplification to be stalled when using native archaeal DNA polymerases. This is a serious setback since the repetitive heating and cooling cycles of a DNA sample being amplified by PCR results in partial, thermally induced deamination of dCTP (a component incorporated into newly amplified DNA) to dUTP (which can be incorporated into DNA) and deamination of deoxycytidine in the DNA to deoxyuracil. This can result in the native archaeal DNA polymerases being unsuitable for PCRs, in particular those concerned with the prevention of "carry-over contamination" where PCR is carried out with dUTP rather than dTTP.

It is an object of the present invention to provide modified archaeal DNA polymerases that do not have the disadvantage of being inhibited by deoxyuracil and are particularly useful in polymerase chain reactions.

Accordingly, a first aspect of the present invention provides a variant archaeal DNA polymerase having a modified amino acid sequence of a wild-type amino acid sequence, the modified sequence being in the amino-terminal amino acids that comprise a uracil-binding pocket in the wild-type polymerase whereby the variant polymerase has reduced affinity for uracil compared to wild-type archaeal DNA polymerases.

The present invention is based upon research (see the Examples) conducted by the inventors that has identified a uracil-binding pocket in archaeal DNA polymerases. They

realised that this pocket may be altered to provide variant polymerases according to the invention that may be beneficially used as described herein.

The variant archaeal DNA polymerase may be a modification of an archaeal family B DNA polymerase. For instance the variant may be derived from any one of the fourteen archaeal family B DNA polymerase shown in Figure 1. For instance the variant may be derived from the polymerases found in *Pyrococcus furiosus* (Pfu-Pol), *Thermococcus gorgonarius* (Tgo-Pol), *Thermococcus litoralis* (Tli-Pol), *Thermococcus sp.* 9°N-7 (9°N-7-Pol), *Desulfurococcus* strain Tok (DTok-Pol), *Pyrobaculum islandicum* (Pis-Pol), *Archaeoglobus fulgidus* (Afu-Pol), *Sulfolobus acidocaldarius* (Sac-Pol), *Sulfurisphaera ohwakuensis* (Soh-Pol), *Sulfolobus solfataricus* (Sso-Pol), *Pyrodictium occultum* (Poc-Pol) or *Aeropyrum pernix* (Ape-Pol). It will be appreciated that the variant could also be derived from any other archaeal family B DNA polymerase.

It is preferred that the variant is derived from *Pyrococcus furiosus* (Pfu-Pol). Wild-type Pfu-Pol used by the inventors has the following amino acid sequence:

MAILDVDYITEEGKPVIRLFKKENGKFKIEHDRTFRPYIYALLRDDSKIEEVKKITGERHGKIVRIVDVEK
VEKKFLGKPITVWKLYLEHPQDVPTIREKVREHPAVVDIFEYDIPFAKRYLIDKGLIPMEGEEELKILAF
DIETLYHEGEEFGKGPIIMISYADENEAKVITWKNIDLPYVEVVSSEREMIKRFLRIIREKDPDIIVTYN
GDSFDFPYLAKRAEKLGIKLTIGRDGSEPKMQRIGDMTAVEVKGRIHFDLYHVITRTINLPTYTLEAVYE
AIFGKPKEKVYADEIAKAWESGENLERVAKYSMEDAKATYELGKEFLPMEIQLSRLVGQPLWDVSRSSTG
NLVEWFLLRKAYERNEVAPNKPSEEEYQRRLRESYTGGFVKEPEKGLWENIVYLDFRALYPSIIITHNVS
PDTLNLEGCKNYDIAPQVGHKFCKDIPGFIPSLLGHLLEERQKIKTKMKETQDPIEKILLDYRQKAIKLL
ANSFYGYYGYAKARWYCKECAESVTAWGRKYIELVWKELEEKFGFKVLYIDTDGLYATIPGGESEEIKKK
ALEFVKYINSKLPGLLELEYEGFYKRGFFVTKKRYAVIDEEGKVITRGLEIVRRDWSEIAKETQARVLET
ILKHGDVEEAVRIVKEVIQKLANYEIPPEKLAIYEQITRPLHEYKAIGPHVAVAKKLAAKGVKIKPGMVI
GYIVLRGDGPISNRAILAEEYDPKKHKYDAEYYIENQVLPAVLRILEGFGYRKEDLRYQKTRQVGLTSWL
NIKKS

The Pfu Pol used by the inventors contains an extra A (underlined in the sequence above) at position 2. This extra amino acid was incorporated because it improves protein expression without affecting the properties of the enzyme. The true wild type Pfu Pol begins MILDVDY....

The inventors have found (see Example 1) that a uracil-binding pocket of the wild type polymerase forms part of the ssDNA template binding cleft of the polymerase (the so call cleft T). Furthermore the uracil-binding pocket comprises amino acids from two conserved regions of the polymerases: Region A and Region B separated by an unconserved region. In the archaeal polymerase from *Pyrococcus furiosus* (Pfu-Pol) Region A is formed by the amino acids 1-40 and Region B by amino acids 78-130. Highly conserved residues in these two regions form the highly ordered uracil-binding pocket. Other archaea have similar regions A & B in their respective polymerases as illustrated in Figure 1. Preferably, one or more of the amino acids in Regions A and/or B are altered to form the variant archaeal DNA polymerase.

Figure 1 illustrates a sequence alignment of the N -terminal domains of various archaeal polymerases. In figure 1 amino acids designated (1) have 90% or greater identity, (2) indicates 80- 90% identity and (3) 60-80% identity. The two highly conserved regions that form the uracil binding pocket are:

Region A, amino acids 1-40 in Pfu-Pol (and corresponding regions in the other polymerases); and

Region B, amino acids 78-131 in Pfu-Pol (and corresponding regions in the other polymerases).

It is preferred that the variant is formed by alteration of one of the amino acids block shaded (1, 2 or 3) in Figure 1. For example the alternation may be in the motif: E - - I - F/Y- - -Y- -D.

The alteration may consist of a substitution, deletion or addition. One of the invariant residues may be altered or other residues in Regions A and/or B that affect the conformation of the uracil-binding pocket.

The inventors believe that residues 7, 36, 37, 90-97 and 112 – 119 in Pfu-Pol are particularly important for uracil binding. Preferably, at least one of these residues is altered to effect the conformation of the pocket and thereby reduce its uracil-binding ability. More preferably, the mutation in Pfu-Pol consists of a change in the amino acids Y7, Y37, V93 or P115. More preferably, the change consists of Y7A, Y37A, V93Q or P115Δ. A most preferred Pfu-Pol mutation is V93Q.

Examples of preferred Pfu-Pol variants have the following amino acid sequences:

(a) **Pfu pol Y7A (Y8A)**

MAILDVDAITEEGKPVIRLFKKENGKFKIEHDRTFRPYIYALLRDDSKIEEVKKITGERHGKIVRIVDVEK
VEKKFLGKPITVWKLYLEHPQDVPTIREKVREHPAVVDIFEYDIPFAKRYLIDKGLIPMEGEEELKILAF
DIETLYHEGEEFGKGPIIMISYADENEAKVITWKNIDLPYVEVVSSEREMIKRFLRIIREKDPDIIVTYN
GDSFDFPYLAKRAEKLGIKLTIGRDGSEPKMQRIGDMTAVEVKGRIHFDLYHVITRTINLPTYTLEAVYE
AIFGKPKEKVYADEIAKAWESGENLERVAKYSMEDAKATYELGKEFLPMEIQLSRLVGQPLWDVSRSSTG
NLVEWFLLRKAYERNEVAPNKPSEEEYQRRLRESYTGGFVKEPEKGLWENIVYLDFRALYPSIIITHNVS
PDTLNLEGCKNYDIAPQVGHKFCKDIPGFIPSLLGHLLEERQKIKTKMKETQDPIEKILLDYRQKAIKLL
ANSFYGYYGYAKARWYCKECAESVTAWGRKYIELVWKELEEKFGFKVLYIDTDGLYATIPGGESEEIKKK

ALEFVKYINSKLPGLLELEYEGFYKRGFFVTKKRYAVIDEEGKVITRGLEIVRRDWSEIAKETQARVLET
ILKHGDVEEAVRIVKEVIQKLANYEIPPEKLAIYEQITRPLHEYKAIGPHVAVAKKLAAKGVKIKPGMVI
GYIVLRGDGPISNRAILAEEYDPKKHKYDAEYYIENQVLPAVLRILEGFGYRKEDLRYQKTRQVGLTSWL
NIKKS

(b) **Pfu pol Y37A (Y38A)**

MAILDVDYITEEGKPVIRLFKKENGKFKIEHDRTFRPAIYALLRDDSKIEEVKKITGERHGKIVRIVDVEK
VEKKFLGKPITVWKLYLEHPQDVPTIREKVREHPAVVDIFEYDIPFAKRYLIDKGLIPMEGEEELKILAF
DIETLYHEGEEFGKGPIIMISYADENEAKVITWKNIDLPYVEVVSSEREMIKRFLRIIREKDPDIIVTYN
GDSFDFPYLAKRAEKLGIKLTIGRDGSEPKMQRIGDMTAVEVKGRIHFDLYHVITRTINLPTYTLEAVYE
AIFGKPKEKVYADEIAKAWESGENLERVAKYSMEDAKATYELGKEFLPMEIQLSRLVGQPLWDVSRSSTG
NLVEWFLLRKAYERNEVAPNKPSEEEYQRRLRESYTGGFVKEPEKGLWENIVYLDFRALYPSIIITHNVS
PDTLNLEGCKNYDIAPQVGHKFCKDIPGFIPSLLGHLLEERQKIKTKMKETQDPIEKILLDYRQKAIKLL
ANSFYGYYGYAKARWYCKECAESVTAWGRKYIELVWKELEEKFGFKVLYIDTDGLYATIPGGESEEIKKK
ALEFVKYINSKLPGLLELEYEGFYKRGFFVTKKRYAVIDEEGKVITRGLEIVRRDWSEIAKETQARVLET
ILKHGDVEEAVRIVKEVIQKLANYEIPPEKLAIYEQITRPLHEYKAIGPHVAVAKKLAAKGVKIKPGMVI
GYIVLRGDGPISNRAILAEEYDPKKHKYDAEYYIENQVLPAVLRILEGFGYRKEDLRYQKTRQVGLTSWL
NIKKS

(c) **Pfu pol V93Q (V94Q)**

MAILDVDYITEEGKPVIRLFKKENGKFKIEHDRTFRPYIYALLRDDSKIEEVKKITGERHGKIVRIVDVEK
VEKKFLGKPITVWKLYLEHPQDQPTIREKVREHPAVVDIFEYDIPFAKRYLIDKGLIPMEGEEELKILAF
DIETLYHEGEEFGKGPIIMISYADENEAKVITWKNIDLPYVEVVSSEREMIKRFLRIIREKDPDIIVTYN
GDSFDFPYLAKRAEKLGIKLTIGRDGSEPKMQRIGDMTAVEVKGRIHFDLYHVITRTINLPTYTLEAVYE
AIFGKPKEKVYADEIAKAWESGENLERVAKYSMEDAKATYELGKEFLPMEIQLSRLVGQPLWDVSRSSTG
NLVEWFLLRKAYERNEVAPNKPSEEEYQRRLRESYTGGFVKEPEKGLWENIVYLDFRALYPSIIITHNVS
PDTLNLEGCKNYDIAPQVGHKFCKDIPGFIPSLLGHLLEERQKIKTKMKETQDPIEKILLDYRQKAIKLL
ANSFYGYYGYAKARWYCKECAESVTAWGRKYIELVWKELEEKFGFKVLYIDTDGLYATIPGGESEEIKKK
ALEFVKYINSKLPGLLELEYEGFYKRGFFVTKKRYAVIDEEGKVITRGLEIVRRDWSEIAKETQARVLET
ILKHGDVEEAVRIVKEVIQKLANYEIPPEKLAIYEQITRPLHEYKAIGPHVAVAKKLAAKGVKIKPGMVI
GYIVLRGDGPISNRAILAEEYDPKKHKYDAEYYIENQVLPAVLRILEGFGYRKEDLRYQKTRQVGLTSWL
NIKKS

It will be appreciated that the preferred mutants listed above comprise an alanine (A) insertion at position 2 which is not found in the wild type. Accordingly such mutants may be designated Y8A, Y38A and V94Q mutants of the MAILDVDY form of Pfu Pol and correspond to Y7A, Y37A and V93Q mutants of the true wild-type (MILDVDY). Y7A, Y37A and V93Q mutants of the true wild-type (MILDVDY) are also preferred mutants according to the invention.

It will be appreciated that equivalent residues in other archaeal polymerases may be mutated (see figure 1). For instance Y7A, Y37A, V93Q and P115Δ remain preferred mutants in Tgo-pol, DTok-pol and 9°N-7-pol.

The present invention also provides nucleic acids encoding an archaeal DNA polymerase variant as defined above.

Preferred mutant genes have the following DNA sequences:

(a) **Pfu pol Y7A (Y8A)**

ATGGCTATCCTGGACGTTGACGCCATCACCGAAGAAGGTAAGCCGGTTATCCGTCTGTTCAAAAAAGAAAACGGT
AAATTCAAAATCGAACACGACCGTACCTTCCGTCCGTACATCTACGCTCTGCTGCGTGACGACTCTAAAATCGAA
GAAGTTAAAAAAATCACCGGTGAACGTCATGGAAAGATTGTGAGAATTGTTGATGTAGAGAAGGTTGAGAAAAAG
TTTCTCGGCAAGCCTATTACCGTGTGGAAACTTTATTTGGAACATCCCCAAGATGTTCCCACTATTAGAGAAAAA
GTTAGAGAACATCCAGCAGTTGTGGACATCTTCGAATACGATATTCCATTTGCAAAGAGATACCTCATCGACAAA
GGCCTAATACCAATGGAGGGGGAAGAAGAGCTAAAGATTCTTGCCTTCGATATAGAAACCCTCTATCACGAAGGA
GAAGAGTTTGGAAAAGGCCCAATTATAATGATTAGTTATGCAGATGAAAATGAAGCAAAGGTGATTACTTGGAAA
AACATAGATCTTCCATACGTTGAGGTTGTATCAAGCGAGAGAGAGATGATAAAGAGATTTCTCAGGATTATCAGG
GAGAAGGATCCTGACATTATAGTTACTTATAATGGAGACTCATTCGACTTCCCATATTTAGCGAAAAGGGCAGAA
AAACTTGGGATTAAATTAACCATTGGAAGAGATGGAAGCGAGCCCAAGATGCAGAGAATAGGCGATATGACGGCT
GTAGAAGTCAAGGGAAGAATACATTTCGACTTGTATCATGTAATAACAAGGACAATAAATCTCCCAACATACACA
CTAGAGGCTGTATATGAAGCAATTTTTGGAAAGCCAAAGGAGAAGGTATACGCCGACGAGATAGCAAAAGCCTGG
GAAAGTGGAGAGAACCTTGAGAGAGTTGCCAAATACTCGATGGAAGATGCAAAGGCAACTTATGAACTCGGGAAA
GAATTCCTTCCAATGGAAATTCAGCTTTCAAGATTAGTTGGACAACCTTTATGGGATGTTTCAAGGTCAAGCACA
GGGAACCTTGTAGAGTGGTTCTTACTTAGGAAAGCCTACGAAAGAAACGAAGTAGCTCCAAACAAGCCAAGTGAA
GAGGAGTATCAAAGAAGGCTCAGGGAGAGCTACACAGGTGGATTCGTTAAAGAGCCAGAAAAGGGGTTGTGGGAA
AACATAGTATACCTAGATTTTAGAGCCCTATATCCCTCGATTATAATTACCCACAATGTTTCTCCCGATACTCTA
AATCTTGAGGGATGCAAGAACTATGATATCGCTCCTCAAGTAGGCCACAAGTTCTGCAAGGACATCCCTGGTTTT
ATACCAAGTCTCTTGGGACATTTGTTAGAGGAAAGACAAAAGATTAAGACAAAAATGAAGGAAACTCAAGATCCT
ATAGAAAAAATACTCCTTGACTATAGACAAAAAGCGATAAAACTCTTAGCAAATTCTTTCTACGGATATTATGGC
TATGCAAAAGCAAGATGGTACTGTAAGGAGTGTGCTGAGAGCGTTACTGCCTGGGGAAGAAAGTACATCGAGTTA
GTATGGAAGGAGCTCGAAGAAAAGTTTGGATTTAAAGTCCTCTACATTGACACTGATGGTCTCTATGCAACTATC
CCAGGAGGAGAAAGTGAGGAAATAAAGAAAAAGGCTCTAGAATTTGTAAAATACATAAATTCAAAGCTCCCTGGA
CTGCTAGAGCTTGAATATGAAGGGTTTTATAAGAGGGGATTCTTCGTTACGAAGAAGAGGTATGCAGTAATAGAT
GAAGAAGGAAAAGTCATTACTCGTGGTTTAGAGATAGTTAGGAGAGATTGGAGTGAAATTGCAAAAGAAACTCAA
GCTAGAGTTTTGGAGACAATACTAAAACACGGAGATGTTGAAGAAGCTGTGAGAATAGTAAAAGAAGTAATACAA
AAGCTTGCCAATTATGAAATTCCACCAGAGAAGCTCGCAATATATGAGCAGATAACAAGACCATTACATGAGTAT
AAGGCGATAGGTCCTCACGTAGCTGTTGCAAAGAAACTAGCTGCTAAAGGAGTTAAAATAAAGCCAGGAATGGTA
ATTGGATACATAGTACTTAGAGGCGATGGTCCAATTAGCAATAGGGCAATTCTAGCTGAGGAATACGATCCCAAA
AAGCACAAGTATGACGCAGAATATTACATTGAGAACCAGGTTCTTCCAGCGGTACTTAGGATATTGGAGGGATTT
GGATACAGAAAGGAAGACCTCAGATACCAAAAGACAAGACAAGTCGGCCTAACTTCCTGGCTTAACATTAAAAAA
TCC

(b) **Pfu pol V93Q (V94Q)**

ATGGCTATCCTGGACGTTGACTACATCACCGAAGAAGGTAAGCCGGTTATCCGTCTGTTCAAAAAAGAAAACGGT
AAATTCAAAATCGAACACGACCGTACCTTCCGTCCGTACATCTACGCTCTGCTGCGTGACGACTCTAAAATCGAA
GAAGTTAAAAAAATCACCGGTGAACGTCATGGAAAGATTGTGAGAATTGTTGATGTAGAGAAGGTTGAGAAAAAG
TTTCTCGGCAAGCCTATTACCGTGTGGAAACTTTATTTGGAACATCCCCAAGATCAGCCCACTATTAGAGAAAAA
GTTAGAGAACATCCAGCAGTTGTGGACATCTTCGAATACGATATTCCATTTGCAAAGAGATACCTCATCGACAAA
GGCCTAATACCAATGGAGGGGGAAGAAGAGCTAAAGATTCTTGCCTTCGATATAGAAACCCTCTATCACGAAGGA
GAAGAGTTTGGAAAAGGCCCAATTATAATGATTAGTTATGCAGATGAAAATGAAGCAAAGGTGATTACTTGGAAA
AACATAGATCTTCCATACGTTGAGGTTGTATCAAGCGAGAGAGAGATGATAAAGAGATTTCTCAGGATTATCAGG
GAGAAGGATCCTGACATTATAGTTACTTATAATGGAGACTCATTCGACTTCCCATATTTAGCGAAAAGGGCAGAA
AAACTTGGGATTAAATTAACCATTGGAAGAGATGGAAGCGAGCCCAAGATGCAGAGAATAGGCGATATGACGGCT
GTAGAAGTCAAGGGAAGAATACATTTCGACTTGTATCATGTAATAACAAGGACAATAAATCTCCCAACATACACA
CTAGAGGCTGTATATGAAGCAATTTTTGGAAAGCCAAAGGAGAAGGTATACGCCGACGAGATAGCAAAAGCCTGG
GAAAGTGGAGAGAACCTTGAGAGAGTTGCCAAATACTCGATGGAAGATGCAAAGGCAACTTATGAACTCGGGAAA
GAATTCCTTCCAATGGAAATTCAGCTTTCAAGATTAGTTGGACAACCTTTATGGGATGTTTCAAGGTCAAGCACA
GGGAACCTTGTAGAGTGGTTCTTACTTAGGAAAGCCTACGAAAGAAACGAAGTAGCTCCAAACAAGCCAAGTGAA
GAGGAGTATCAAAGAAGGCTCAGGGAGAGCTACACAGGTGGATTCGTTAAAGAGCCAGAAAAGGGGTTGTGGGAA
AACATAGTATACCTAGATTTTAGAGCCCTATATCCCTCGATTATAATTACCCACAATGTTTCTCCCGATACTCTA
AATCTTGAGGGATGCAAGAACTATGATATCGCTCCTCAAGTAGGCCACAAGTTCTGCAAGGACATCCCTGGTTTT
ATACCAAGTCTCTTGGGACATTTGTTAGAGGAAAGACAAAAGATTAAGACAAAAATGAAGGAAACTCAAGATCCT
ATAGAAAAAATACTCCTTGACTATAGACAAAAAGCGATAAAACTCTTAGCAAATTCTTTCTACGGATATTATGGC
TATGCAAAAGCAAGATGGTACTGTAAGGAGTGTGCTGAGAGCGTTACTGCCTGGGGAAGAAAGTACATCGAGTTA
GTATGGAAGGAGCTCGAAGAAAAGTTTGGATTTAAAGTCCTCTACATTGACACTGATGGTCTCTATGCAACTATC
CCAGGAGGAGAAAGTGAGGAAATAAAGAAAAAGGCTCTAGAATTTGTAAAATACATAAATTCAAAGCTCCCTGGA
CTGCTAGAGCTTGAATATGAAGGGTTTTATAAGAGGGGATTCTTCGTTACGAAGAAGAGGTATGCAGTAATAGAT
GAAGAAGGAAAAGTCATTACTCGTGGTTTAGAGATAGTTAGGAGAGATTGGAGTGAAATTGCAAAAGAAACTCAA
GCTAGAGTTTTGGAGACAATACTAAAACACGGAGATGTTGAAGAAGCTGTGAGAATAGTAAAAGAAGTAATACAA
AAGCTTGCCAATTATGAAATTCCACCAGAGAAGCTCGCAATATATGAGCAGATAACAAGACCATTACATGAGTAT
AAGGCGATAGGTCCTCACGTAGCTGTTGCAAAGAAACTAGCTGCTAAAGGAGTTAAAATAAAGCCAGGAATGGTA
ATTGGATACATAGTACTTAGAGGCGATGGTCCAATTAGCAATAGGGCAATTCTAGCTGAGGAATACGATCCCAAA
AAGCACAAGTATGACGCAGAATATTACATTGAGAACCAGGTTCTTCCAGCGGTACTTAGGATATTGGAGGGATTT
GGATACAGAAAGGAAGACCTCAGATACCAAAAGACAAGACAAGTCGGCCTAACTTCCTGGCTTAACATTAAAAAA
TCC

(b) **Pfu Pol P115Δ (P116Δ)**

ATGGCTATCCTGGACGTTGACTACATCACCGAAGAAGGTAAGCCGGTTATCCGTCTGTTCAAAAAAGAAAACGGT
AAATTCAAAATCGAACACGACCGTACCTTCCGTCCGTACATCTACGCTCTGCTGCGTGACGACTCTAAAATCGAA
GAAGTTAAAAAAATCACCGGTGAACGTCATGGAAAGATTGTGAGAATTGTTGATGTAGAGAAGGTTGAGAAAAAG
TTTCTCGGCAAGCCTATTACCGTGTGGAAACTTTATTTGGAACATCCCCAAGATGTTCCCACTATTAGAGAAAAA
GTTAGAGAACATCCAGCAGTTGTGGACATCTTCGAATACGATATTTTTGCAAAGAGATACCTCATCGACAAAGGC
CTAATACCAATGGAGGGGGAAGAAGAGCTAAAGATTCTTGCCTTCGATATAGAAACCCTCTATCACGAAGGAGAA
GAGTTTGGAAAAGGCCCAATTATAATGATTAGTTATGCAGATGAAAATGAAGCAAAGGTGATTACTTGGAAAAAC
ATAGATCTTCCATACGTTGAGGTTGTATCAAGCGAGAGAGAGATGATAAAGAGATTTCTCAGGATTATCAGGGAG
AAGGATCCTGACATTATAGTTACTTATAATGGAGACTCATTCGACTTCCCATATTTAGCGAAAAGGGCAGAAAAA
CTTGGGATTAAATTAACCATTGGAAGAGATGGAAGCGAGCCCAAGATGCAGAGAATAGGCGATATGACGGCTGTA
GAAGTCAAGGGAAGAATACATTTCGACTTGTATCATGTAATAACAAGGACAATAAATCTCCCAACATACACACTA
GAGGCTGTATATGAAGCAATTTTTGGAAAGCCAAAGGAGAAGGTATACGCCGACGAGATAGCAAAAGCCTGGGAA
AGTGGAGAGAACCTTGAGAGAGTTGCCAAATACTCGATGGAAGATGCAAAGGCAACTTATGAACTCGGGAAAGAA
TTCCTTCCAATGGAAATTCAGCTTTCAAGATTAGTTGGACAACCTTTATGGGATGTTTCAAGGTCAAGCACAGGG
AACCTTGTAGAGTGGTTCTTACTTAGGAAAGCCTACGAAAGAAACGAAGTAGCTCCAAACAAGCCAAGTGAAGAG
GAGTATCAAAGAAGGCTCAGGGAGAGCTACACAGGTGGATTCGTTAAAGAGCCAGAAAAGGGGTTGTGGGAAAAC
ATAGTATACCTAGATTTTAGAGCCCTATATCCCTCGATTATAATTACCCACAATGTTTCTCCCGATACTCTAAAT

CTTGAGGGATGCAAGAACTATGATATCGCTCCTCAAGTAGGCCACAAGTTCTGCAAGGACATCCCTGGTTTTATA
CCAAGTCTCTTGGGACATTTGTTAGAGGAAAGACAAAAGATTAAGACAAAAATGAAGGAAACTCAAGATCCTATA
GAAAAAATACTCCTTGACTATAGACAAAAAGCGATAAAACTCTTAGCAAATTCTTTCTACGGATATTATGGCTAT
GCAAAAGCAAGATGGTACTGTAAGGAGTGTGCTGAGAGCGTTACTGCCTGGGGAAGAAAGTACATCGAGTTAGTA
TGGAAGGAGCTCGAAGAAAAGTTTGGATTTAAAGTCCTCTACATTGACACTGATGGTCTCTATGCAACTATCCCA
GGAGGAGAAAGTGAGGAAATAAAGAAAAAGGCTCTAGAATTTGTAAAATACATAAATTCAAAGCTCCCTGGACTG
CTAGAGCTTGAATATGAAGGGTTTTATAAGAGGGGATTCTTCGTTACGAAGAAGAGGTATGCAGTAATAGATGAA
GAAGGAAAAGTCATTACTCGTGGTTTAGAGATAGTTAGGAGAGATTGGAGTGAAATTGCAAAAGAAACTCAAGCT
AGAGTTTTGGAGACAATACTAAAACACGGAGATGTTGAAGAAGCTGTGAGAATAGTAAAAGAAGTAATACAAAAG
CTTGCCAATTATGAAATTCCACCAGAGAAGCTCGCAATATATGAGCAGATAACAAGACCATTACATGAGTATAAG
GCGATAGGTCCTCACGTAGCTGTTGCAAAGAAACTAGCTGCTAAAGGAGTTAAAATAAAGCCAGGAATGGTAATT
GGATACATAGTACTTAGAGGCGATGGTCCAATTAGCAATAGGGCAATTCTAGCTGAGGAATACGATCCCAAAAAG
CACAAGTATGACGCAGAATATTACATTGAGAACCAGGTTCTTCCAGCGGTACTTAGGATATTGGAGGGATTTGGA
TACAGAAAGGAAGACCTCAGATACCAAAAGACAAGACAAGTCGGCCTAACTTCCTGGCTTAACATTAAAAAATCC

It will be appreciated that preferred mutants encoded by the DNA sequences listed above comprise a codon for an alanine (A) inserted at position 2 that is not found in the wild type. Accordingly such mutants may be designated Y8A, Y38A and P116Δ mutants of the MAILDVDY form of Pfu Pol and correspond to Y7A, Y37A and P115Δ mutants of the true wild-type (MILDVDY). Nucleic acid molecules encoding Y7A, Y37A and P115Δ mutants of the true wild-type (MILDVDY) are also preferred nucleic acids according to the invention.

The variant polymerases as defined above are particularly useful for PCRs since they are thermally stable, have proof-reading ability but are not stalled by the presence of dUTP.

Accordingly, a further aspect of the present invention provides a kit useful for polymerase chain reactions comprising DNA to be amplified, free bases, primers and a variant archaeal DNA polymerase having a modified amino acid sequence of a wild-type amino acid sequence, the modified sequence being in the amino-terminal amino acids that comprise a uracil binding pocket in the wild-type polymerase whereby the variant polymerase has reduced affinity for uracil compared to the wild-type polymerase.

The present invention further provides a method of amplifying DNA comprising the steps of (i) denaturing a double strand of DNA by heating a solution containing the DNA, free oligonucleotides, primers and a variant archaeal DNA polymerase having a modified amino acid sequence of a wild-type amino acid sequence, the modified sequence being in the amino-terminal amino acids that comprise a uracil binding pocket in the wild type polymerase whereby the variant polymerase has reduced affinity for uracil compared to the wild-type polymerase; (ii) reducing the temperature of the solution to effect annealing of the primer and the DNA and (iii) heating the solution to effect extension of DNA by the variant polymerase.

The ability of the variant DNA polymerase to amplify DNA in the presence of dUTP results in it being particularly suitable for PCR that uses dUTP rather than dTTP, for example in the prevention of contamination of samples.

A preferred protocol for carrying out PCR utlising variant polymerases according to the method of the present invention is as follows: PCR may carried out under the following conditions: 100µl volume, 20 mM Tris-HCl pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X100, 100 µg/ml BSA, 250 µM each of dATP, dGTP, dCTP and either 250 µM dTTP or 250 µM dUTP, 2.5 units DNA polymerase overlayed with 40µl of mineral oil (1 unit of polymerase is defined as amount of enzyme that incorporates 10 nmol of dATP into acid-precipitable material using an activated calf-thymus DNA-based assay (4) in 30 min at 72°C). 5 ng of template DNA to be amplified is used and the concentration of the forward

and reverse primers (each 18 bases in length) is 0.3 μM. Each PCR consisted of 30 cycles of 1 min at 95°C, 2 min at 52°C and 4.5 min at 72°C.

The present invention will now be further illustrated by means of the following Examples in which Example 1 investigates the uracil-binding pocket of Archaea polymerases, in particular the hyperthermophilic archaea, *Pyrococcus furiosus* (Pfu-pol) and Example 2 investigates the effect of mutagenesis of residues in and around the uracil-binding pocket of archaeal DNA polymerases, and with reference to the accompanying drawings in which:-

**Figure 1** illustrates an amino acid sequence alignment of archael family B DNA polymerises, corresponding to residues 1-130 of *Pyrococcus furiosus* polymerase. Candidates were identified using a WUBLAST search (European Bioinformatics Institute, http://www.ebi.ac.uk/ebi_home.html) for homologues of Pfu-Pol. An additional ENTREZ search of the SWISSPROT database for family B DNA polymerises was performed [Genbank (http://www.ncbi.nlm.nih.gov/)]. Sequence alignments were generated using ClustalX (version 1.81) [J. D. Thompson *et al., Nucl. Acids. Res* 24, 4876 (1997]. The organisms and the DNA polymerase sequence accession numbers were: *Pyrococcus fuiosus* (Pfu) (P80061), *Thermococcus gorgonarius* (Tgo) (pdb 1D5A), *Pyrococcus kodakaraensis* (PKOD) (gi/13399597), *Desulfurococcus* strain Tok (DTok), *Thermococcus sp.* 9°N-7 (9°N-7) (Q56366), *Thermococcus litoralis* (Tli) (AAA72101.1), *Methanococcus voltae* (Mvo) (P52025), *Pyrobaculum islandicum* (Pis) (AAF27815.1), *Archaeoglobus fulgidus* (Agu) (O29753), *Cenarchaeaum symbiosum* (Csy) (AAC62712.1), *Sulfolobus acidocaldarius* (Sac) (P95690), *Sulfurisphaera ohwakuensis* (Soh) (BAA23994.1), *Sulfolobus solfataricus* (Sso)

(P26811), *Pyrodictium occultum* (Poc) (BAA07579.1) and *Aeropyrum pernix* (Ape) (NP_148473.1);

**Figure 2A:** The template-binding cleft T of Tgo-Pol showing the presence of a pocket. **B:** The N-terminal domain of Tgo-Pol with amino acids that form the pocket shown in space-fill: Y7; P36/Y37; amino acids 90-97; amino acids 112-116. α-Helices are shown (4) and β-sheets (5). **C:** Amino acid sequences of the N-terminal domains of Tgo-Pol (upper sequence) and RB69-Pol (lower sequence). Amino acids that form the pocket in Tgo-Pol (and the corresponding residues in RB69-Pol) are underlined and correspond to the amino acids identified in panels B and E. Cylinders represent α-helices and arrows β-sheets. The amino acid sequences have minimal homology and have been aligned using structural homology. **D:** Structural alignment of the N-terminal domains of Tgo-Pol and RB69-Pol. The insert in Tgo-Pol is shown. This structural superimposition was used to generate the amino acid alignment shown in C. **E:** The N-terminal domain of RB69-Pol. Space-filled amino acids (V8 Q10; residues 65-72; residues 84-89; P35/S36 lie substantially behind residues 84-89) correspond to those in Tgo-Pol (shown in panel B) which form the pocket. Both V8 and Q10 are near the position of the Tgo-Pol Y7. α-Helices are shown (4) and β-sheets (5). Images/structural homology models were produced using Swiss-Model [N. Guex, M. C. Peitsch, *Electrophoresis* **18**, 2714 (1997)] (http://www.expasy.ch/spdbv/), POV-Ray [C. Cason, POV-Ray for Windows, version 3.1g (1999)] (http://www.povray.org) and Rasmol [R. Sayle, J. F. Milner-White, *Trends Biochem. Sci.* **20**, 374 (1995)].

**Figure 3.** Modelling pyrimidines into the N-terminal domain pocket of Tgo-Pol using Web Lab Viewer Pro [Molecular Simulations Inc., Web Lab Viewer Pro (version 4.0) (2000)]

(http://www.msi.com). Hatched lines represent hydrogen bonds; except steric clashes are identified (6). Uracil formed four enzyme-base hydrogen bonds and no clashes. Cytosine and thymine resulted in fewer hydrogen bonds and/or clashes.

**Figure 4 A and B:** Primer extension reactions using a 5'-$^{32}$P labelled 24-mer primer (5'-GGGGATCCTCTAGAGTCGACCTGC-3') (2.5nM) annealed to 44-mer template (5'-GGAGACAAGCTTG(U/T)ATGCCTGCAGGTCGACTCTAGCGGCTAAAA-3') (5nM). The primer hybridises with the underlined section of the template to place a uracil (thymine in controls) seven bases from the primer-template junction (H.H. Hogrefe, C. J. Hansen, B. R. Scott, K. B. Nielson, *Proc. Natl. Acad. Sci. U.S.A.*, 99, 596 (2002)). In the presence of the four dNTPs all Pfu-Pol variants were able to fully copy the control template, lacking uracil, converting the labelled 24-mer primer into a product 44 bases long (A). With uracil in the template the wild-type enzyme (WT) stalled polymerisation (H.H. Hogrefe, C. J. Hansen, B. R. Scott, K. B. Nielson, *Proc. Natl. Acad. Sci. U.S.A.*, 99, 596 (2002)) giving a truncated product (B). The arrow indicates the major pause site, shown (using standards, not illustrated) to occur four base before uracil. With Y7A, V93Q and P115Δ some full-length product was observed (amounts: V93Q > P115Δ > Y7A) indicating read-through of the template-strand uracil (panel B). No full-length product was seen with Y37A, Y37F and P115F (some material partially extended past the "uracil induced pause" was seen with Y7A, Y37A and Y37F. The first and second lanes contain standard 24-mer and 44-mer. **C and D:** PCR reactions. In control PCR with the four normal dNTPs (dATP, dGTP, dCTP and dTTP) all the Pfu-Pol variants (and Taq-Pol) gave the anticipated PCR product, ~ 2 kbases in length (C). When dUTP replaced dTTP, Taq-Pol produced a 2 kbase fragment (D). With most Pfu-

Pol variants (WT, Y37A, Y37F and P115F) no PCR product was seen (D). Three Pfu-Pol mutant's produced a PCR product (amounts: V93Q > P115Δ > Y7A) (2 x and 4 x loadings, respectively, were used for visualisation with P115Δ and Y7A). All PCR reactions were carried out under identical conditions with no optimisation; this probably accounts for the lower weight contaminant seen in some lanes. Markers (important sizes indicated) were from Promega.

**Example 1**

Crystal structures are known for five family B DNA polymerases; one viral, the remaining four archaeal. The first structure to be solved was for the bacteriophage RB69 polymerase (RB69-Pol) (J. Wang *et al., Cell* 89, 1087 (1997)); a structure with primer-template has also been determined (M. C. Franklin, J. J. Wang, T. A. Steitz, *Cell* **105**, 657 (2001)). More recently, the structure of an archaeal family B DNA polymerase, from the hyperthermophilic archaeon *Thermococcus gorgonarius* (Tgo-Pol), has been reported (K. P. Hopfner *et al., Structure* 7, 1189 (1999)). Three other archaeal polymerase structures, *Desulfurococcus* strain Tok (DTok-Pol) (Y. Zhao *et al., Structure* 7, 1189 (1999)), *Thermococcus sp.* 9°N-7 (9°N-7-Pol) (A. C. Rodgriguez, H-W. Park, C. Mao, L. S. Beese, *J. Mol. Biol.* **299**, 447 (200)) and *Pyrococcus kodakaraensis* KOD1 (KOD1-Pol) (H. Hashimoto *et al., J. Mol. Biol.* **306**, 469 (2001)), were subsequently solved. Only apo-enzyme structures are known for the archaea. All five family B polymerases contain five distinct domains, the N-terminal domain, the exonuclease or 'editing' domain and three polymerase active site domains. The folding of the five domains forms three distinct clefts extending from a central hole. Two (named clefts D and T) are oriented approximately 180° relative to each other, on either side of the central hole. The structure of RB69-pol containing a primer-template demonstrates that cleft D binds double stranded primer-template and cleft T binds single-stranded template (M. C. Franklin *et al supra*). The three polymerase domains forms cleft D, whereas cleft T is formed by the exonuclease domain and the N-terminal domain. The third cleft is perpendicular to the other two and represents the 3'-5' exonuclease/editing cleft.

Examination of cleft T in the case of Tgo-Pol (the other archaeal polymerases give similar results) revealed the presence of a pocket located on a surface exposed face towards the outer edge of the polymerase (figure 2A). The location of this pocket, in the template strand binding region, approximately four bases from the primer-template junction makes it a clear candidate for uracil recognition. The pocket is comprised of amino acids that are solely present in the N-terminal domain. Amino acids from four regions of this domain, which are close together in space, are used to assemble the putative uracil binding pocket. These amino acids are illustrated in figure 2B (a structural representation of the N-terminal domain) and underlined in figure 2C (the amino acid sequence of the N-terminal domain). The function of Y7, which sits at the entrance of the pocket is obscure. It may form a lid, which closes following binding, ensuring trapping (and hence high affinity) of uracil. The base of the pocket (clearly visible as P36/Y37 in figure 2B) is formed by Y37, oriented by P36 at the beginning of a β sheet and supported by K84 "underneath" Y37. One side of the pocket is formed by amino acids 90-97, present in an α-helix. A proline (P94) bends the α-helix, forming a curved wall (figure 1B). The other side comprises residues 110-116; amino acids present in a loop region (110-114) or at the beginning of a second α-helix (115-116). This α-helix commences with amino acids P115 and F116. The proline appears to have a critical role, ensuring that P115 and F116 are able to form part of the pocket's curved wall (figure 2B).

The viral polymerase from RB-69 shows 61% amino acid identity to bacteriophage T4-Pol. Additionally, the N-terminal of RB69-Pol is identical to a structure for an N-terminal fragment of T4-Pol (J. Wang, P. Yu, T. C. Lin, W. H. Konigsberg, T.A. Steitz, *Biochemistry* 35, 8110 (1996)). Previously, it has been demonstrated that T4-Pol did not stall

polymerisation in response to template-strand uracil (M. A. Greagg *et al.*, *Proc Natl. Acad. Sci. U.S.A.* 96, (1999)); inability to recognise uracil would also be expected for RB69-Pol. This contrasts with Tgo-Pol which stalls polymerization when template strand uracil is encountered (M. A. Greagg *et al.*, *Supra*). Tgo-Pol is also inhibited by uracil containing DNA (R. S. Lasken, D. M. Schuster, A. Rashtchian, *J. Biol. Chem.* 271, 17692 (1996)); a feature not seen with viral enzymes. Therefore, if the pocket seen with Tgo-Pol is used for uracil detection, it should be absent for the viral enzymes. The N-terminal domains of Tgo-Pol and RB69-Pol show considerable structural homology (figure 2D); near perfect alignment of secondary structural elements is observed, despite the almost complete lack of amino acid sequence homology (figure 2C). The only significant difference is the presence of an amino acid insert (shown in figure 2D) in the archaeal enzyme. However, detailed comparisons demonstrate differences in the two polymerases (figure 2E). With Tgo-Pol, Y7 may act as a lid and Y36 forms the base of the pocket; in RB-69 these residues are replaced with valine/glutamine (both these amino acids in RB69-Pol are near Y7 of Tgo-Pol and it is not clear which is the exact replacement) and serine respectively. Similarly Tgo-Pol uses a proline-containing kinked α-helix (residues 90-97) to form one wall of the pocket. In RB69-Pol the corresponding α-helix (residues 65-74) lacks the proline; therefore, the helix is straight and does not form a good wall. However, a key feature involves P115 and F116, which form part of one of the walls of the pocket with Tgo-Pol. In the case of RB69 the proline is missing and this results in the corresponding phenylalanine (F88) falling into and completely filling the pocket. As shown in figure 1E this means that the viral enzyme lacks a uracil binding pocket (P35/S36 are substantially obscured). The differences between the viral

and archaeal enzymes, based on subtle changes to a few amino acids, provide compelling evidence that the N-terminal pocket is responsible for uracil detection.

It was possible to model uracil into the pocket of Tgo-Pol (figure 3A). The most favourable orientation produces four hydrogen bonds between the protein and uracil. In all cases the protein uses the peptide backbone for hydrogen bond formation. The interactions comprise: I114 (peptide -NH) to uracil C2 =O group; E111 (peptide =O) to uracil N3H; Y37 (peptide =O) to uracil N3H and (peptide -NH) to uracil C4=O (figure 3A). Cytosine superimposed at the same position as uracil forms only one hydrogen bond (I114 (peptide -NH) to uracil C2 =O) and the 4-$NH_2$ group clashes with the main chain atoms of Y37 (peptide =O and -NH) (figure 3B). This clash could be relieved by repositioning the cytosine, but only at the expense of the one H-bond, resulting in no interactions between the base and the protein. Thymine superimposed at the uracil position could form most of the protein-base hydrogen bonds (interactions were identical to uracil except the Y37 (peptide -NH) to uracil/thymine N3H hydrogen bond was not formed). Critically the C5 -$CH_3$ group showed a severe steric clash with the edges of the cyclic P36 side chain and the ring of F116, thereby preventing binding of the base within the pocket (figure 3C). Thus the pocket is highly specific for binding uracil and able to discriminate against the 'normal' DNA pyrimidines.

An amino acid sequence alignment has been carried out for the N-terminal domains of fourteen archaeal family B DNA polymerases (figure 1); eight were from the crenarchaea and six from the euryarchaea. Twelve of the polymerases were either thermophiles or hyperthermophiles, one was mesophilic (*Methanococcus voltae* (Mvo)) and one was psychrophilic (*Cenarchaeaum symbiosum* (Csy)). Two highly conserved regions (A and B) which contain most of the amino acids that from the uracil binding pocket, are seen. Many of

the amino acids comprising the uracil binding pocket (figure 2B and 2C) are highly conserved; especially uracil-contacting residues (figure 3A). Thus P36, Y37, E111 and I114 show 100 % identity. The possible pocket lid, Y7, also shows 100 % conservation. Several other key amino acids e.g. V93 (which lines one side of the pocket), P115 and F116 (which line the other side of the pocket) show a high degree of conservation.

Example 2

Use was made of three assays to test the ability of the mutant polymerases to recognise uracil. Primer extension reactions (M. A. Greagg *et al.*, *Supra*) measure the ability of a polymerase to extend a primer through uracil bases in the template strand. As expected both the wild type and the mutant enzymes were able to completely copy a control template, lacking uracil (figure 4A). As previously observed (M. A. Greagg *et al.*, *Supra*), the wild type enzyme stalled polymerisation four bases upstream of template-strand uracil, resulting in a truncated product (figure 4B). Y37A, Y37F and P115F behaved in a similar manner to wild type. However, three of the mutant enzymes, V93Q, P115Δ and Y7A, produced full-length product when uracil was present (figure 4B). With V93Q full-length product predominated; in the cases of P115Δ and Y7A both full-length and truncated product were seen.

Next the ability of the polymerases to bind single stranded DNA containing uracil was investigated, using a binding assay based on fluorescence anisotropy (S. L. Reid, D. Parry, H-H. Liu, B. A. Connolly, *Biochemistry* 40, 2484 (2001)). $K_D$ was determined by fluorescence anisotropy used an oligodeoxynucleotide containing a single uracil and a hexachlorofluorescein label at its 5'-terminal. The oligodeoxynucleotide used was: 5'-hex-GCCCGCGGGAUATCGGCCCTTA-3' (or a control in which the uracil was replaced with thymine). The concentration of the oligodeoxynucleotide was 5 nM in 1 ml of 10 mM

Hepes-NaOH, pH7.5, 100mM NaCl, 1mM EDTA. Aliquots of the enzyme were added and the anisotropy measured; titration was continued until the anisotropy stopped increasing. Data fitting to obtain $K_D$ values is as described by Reid *et al* (*Supra*)). $K_D$ values are summarised in table 1. The wild type enzyme bound the uracil-containing oligodeoxynucleotide with a $K_D$ of 8.3 nM, a 17-fold preference over a control strand lacking this base. Three of the mutants, Y7A, V93Q and P115Δ bound to the uracil-containing oligodeoxynucleotide less well than the wild type ($K_D$ values of 25.7, 144.5 and 84 nM respectively, table 1). These mutants correspond exactly to those able to read through a uracil-containing template; furthermore the diminution in uracil binding corresponds with read through capability. In both assays loss of uracil recognition is: V93Q > P115Δ >Y7A. These three mutants also show a reduced preference for uracil-containing DNA over the control sequence; with P115Δ and V93Q the preference is virtually abolished. The mutants Y7A, Y37A and P115F bind uracil-containing DNA with essentially the same affinity as the wild type (table 1). In some cases, Y37A and Y37F, the preference for the uracil-containing oligodeoxynucleotide is reduced, but this arises solely from tighter binding of the control. Table 1 also gives the specific activity of the mutant Pfu-Pols relative to the wild type. In general there are only small decreases in activity with even the mutant with the lowest activity (Y37A) retaining 38 % of the wild type activity.

The DNA polymerase activity assay (Richardson, C. C. (1966) in *Procedures in Nucleic Acids Research*, G. L. Cantoni, D. R. Davies, Eds, (Harper and Row, New York, 1966), pp. 263-27) used 50 µl samples containing 20 mM Tris-HCl, pH 8.8, 10mM KCl, 10mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X100, 100µg/ml BSA, 200µM each dNTPs, 0.2mg/ml activated calf thymus DNA (AP Biotech), 1µCi 3000Ci/mmol [$\alpha$-$^{32}$P] dATP. Pfu-Pol (the amount varied depending on the activity of the enzyme) was added and a 10 minute

incubation at 72°C was used (reactions were linear over this time). After this period the amount of radioactivity incorporated into acid-precipitable material was determined by scintillation counting. 1 unit of enzyme is defined as amount of enzyme that incorporates 10 nmol of dATP into acid-precipitable material in 30 min at 72°C.

**Table 1.** Pfu-Pol variants: specific activity and ability to bind to uracil containing DNA.

| Pfu-Pol variant | specific activity (units/mg) | % activity (relative to wild type) | $K_D$ (nM) (uracil) | $K_D$ (nM) (control) | Preference for uracil |
|---|---|---|---|---|---|
| wild-type | 3556 | 100 | 8 ± 1 | 140 ± 10 | 18 |
| Y7A | 2314 | 65 | 26 ± 1 | 255 ± 20 | 10 |
| Y37A | 1352 | 38 | 9 ± 2 | 57 ±7 | 6 |
| Y37F | 2169 | 61 | 10 ± 4 | 120 ±13 | 12 |
| V93Q | 1643 | 46 | 144 ± 7 | 277 ± 22 | 2 |
| P115F | 2541 | 71 | 7 ± 1 | 148 ± 10 | 21 |
| P115Δ | 1637 | 46 | 84 ± 4 | 111 ± 5 | 1.3 |

The specific activities of the Pfu-Pol variants was determined using the incorporation of [$\alpha$-$^{32}$P]-dATP into acid-precipitable calf-thymus DNA as described (see above). Values are accurate ± 15 %. Binding constants were determined by fluorescence anisotropy using hex-GCCCGCGGGAUATCGGCCCTTA (uracil) or an analogous oligodeoxynucleotide in which the uracil was replaced with thymine (control) (S. L. Reid, *et al*; *supra* and J. Wang *et. al., (Supra)*). Each value was determined three times and the average ± one standard deviation is given. The preference for uracil is the ratio $K_{D\,(uracil)}/K_{D\,(control)}$.

Finally PCR was performed +/- dUTP the ability of Pfu-Pol to carry out PCR was evaluated by amplifying a – 2 kbase fragment (between the T7 promoter and the *Hind*III site) of pET17-b(Pfu-Pol) [S. J. Evans *et al., Nucl. Acids. Res.* 28, 1059 (2000)]. Conditions: 100µl volume, 20 mM Tris-HCl ph 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X100, 100 µg/ml BSA, 250 µM each dNTP (one set of reactions contained dTTP, the other dUTP),

2.5 units DNA polymerase overlayed with 40 µl of mineral oil. 5 ng of pET17-b(Pfu-Pol) was used and the concentrations of the forward and reverse primers were both 0.3 µM. Each reaction was 30 cycles of 1 min at 95°C, 2 min at 52°C and 4.5 min at 72°C. PCR with Taq-Pol was identical to Pfu-Pol save 10 mM Tris-HCl pH 8.8, 50 mM KCl, 0.08% NP-40, 1.5 mM $MgCl_2$ was used. Analysis used ethidium bromide-stained agarose gels.

Pfu-Pol (wild type and mutants) and Taq-Pol, were able to perform PCR with the four normal dNTPs (figure 4C). When dUTP was used in place of dTTP, PCR with Taq-Pol was unaffected; however, wild-type Pfu-Pol gave no product (figure 4D). The three mutants that showed diminished uracil recognition in read-through and binding assays, Y7A, P115Δ and V93Q, gave a PCR product (figure 4D). The amount of PCR product produced was V93Q > P115Δ >Y7A, again matching the order found for loss of uracil recognition. dUTP (concentration 250 µM) completely replaces dTTP in these reactions. The wild type enzyme is completely inhibited when 0.02 µM dUTP is used to spike PCR reactions containing the four normal dNTPs (H. H. Hogrefe, *et al* (*Supra*)); clearly showing these three mutants are very disabled in uracil recognition. The Pfu-Pol mutants, Y37A, Y37F and P115F did not give a PCR product with dUTP.

## CLAIMS

1. A variant archaeal DNA polymerase having a modified amino acid sequence of a wild-type amino acid sequence, the modified sequence being in the amino-terminal amino acids that comprise a uracil-binding pocket in the wild-type polymerase whereby the variant polymerase has reduced affinity for uracil than the wild-type polymerase.

2. A variant archael DNA polymerase according to claim 1 having a modified amino acid sequence of wildtype polymerases selected from *Thermococcus gorgonarius* (Tgo-Pol), *Thermococcus litoralis* (Tli-Pol), *Thermococcus sp.* 9°N-7 (9°N-7-Pol), *Desulfurococcus* strain Tok (DTok-Pol), *Pyrobaculum islandicum* (Pis-Pol), *Archaeoglobus fulgidus* (Afu-Pol), *Sulfolobus acidocaldarius* (Sac-Pol), *Sulfurisphaera ohwakuensis* (Soh-Pol), *Sulfolobus solfataricus* (Sso-Pol), *Pyrodictium occultum* (Poc-Pol) or *Aeropyrum pernix* (Ape-Pol.

3. A variant archaeal DNA polymerase according to claim 1 having a modified amino acid sequence of wildtype *Pyrococcus furiosus* DNA polymerase (Pfu-Pol),

4. A variant archaeal DNA polymerase according to claim 3 having modifications in amino acids 1-40 or amino acids 78-130.

5. A variant archaeal DNA polymerase according to claim 4 having modifications to amino acids 7, 36, 37, 90-97 or 112 – 119.

6. A variant archael DNA polymerase according to claim 5 having modifications to amino acids Y7, V93 or P115.

7. A variant archaeal DNA polymerase according to claim 5 wherein the modification is Y7A.

8. A variant archaeal DNA polymerase according to claim 5 wherein the modification is Y37A.

9. A variant archaeal DNA polymerase according to claim 5 wherein the modification is V93Q.

10. A variant archaeal DNA polymerase according to claim 5 wherein the modification is P115Δ.

11. A variant archael DNA polymerase according to any preceding claim having modifications in the amino acid motif: E - - I -F/Y- - -Y- -D.

12. A nucleic acid molecule encoding an archaeal DNA polymerase according to any one of claims 1 – 11.

13. A method of amplifying DNA comprising the steps of (i) denaturing a double strand of DNA by heating a solution containing the DNA, free oligonucleotides, primers and a

variant archaeal DNA polymerase as defined in any one of claims 1 - 11; (ii) reducing the temperature of the solution to effect annealing of the primer and the DNA and (iii) heating the solution to effect extension of DNA by the variant polymerase.

14. A kit useful for polymerase chain reactions comprising a variant archaeal DNA polymerase as defined in any one of claims 1 - 10 and optionally DNA to be amplified, free bases and primers.

Figure 1

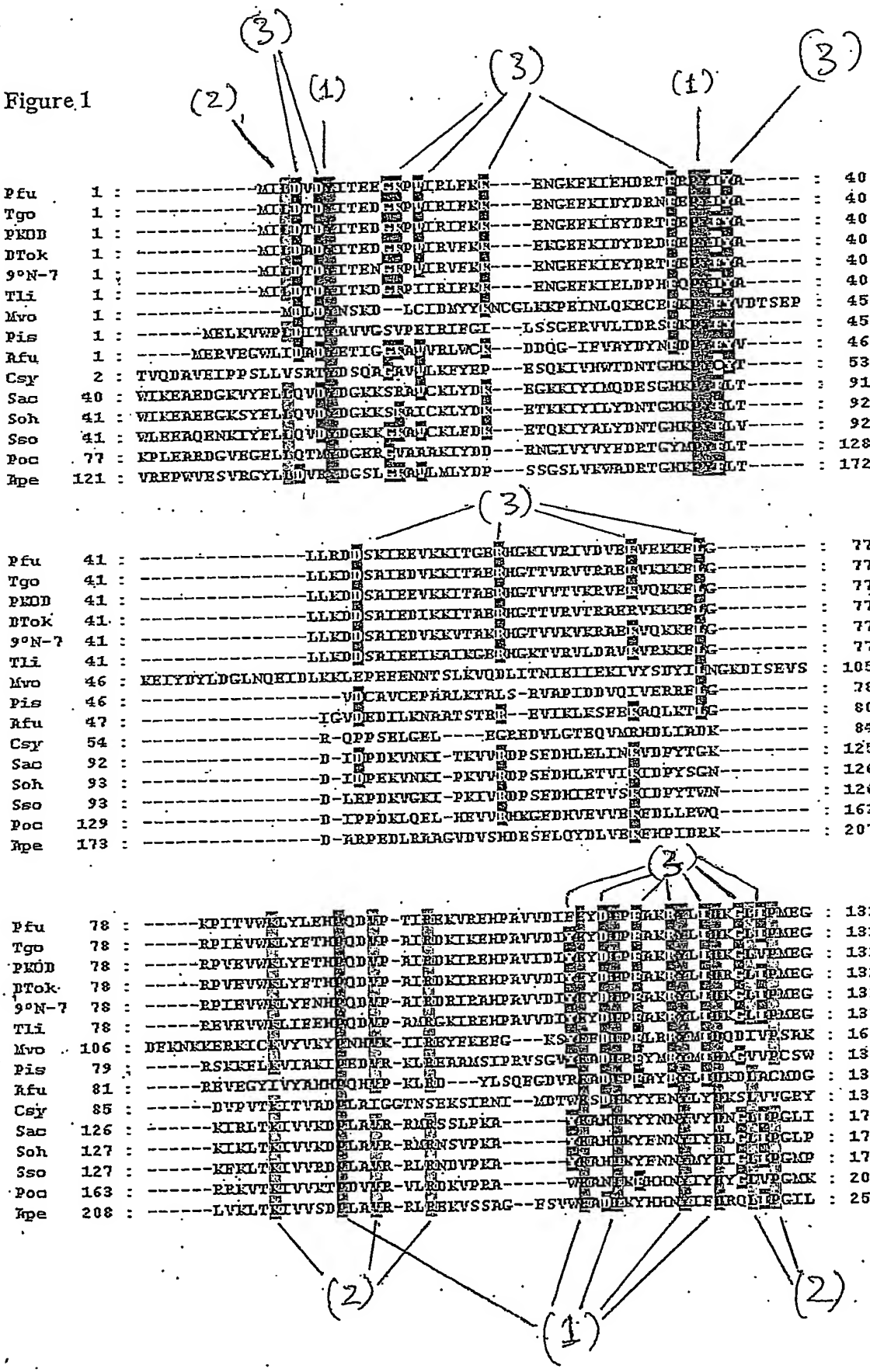

(3)
(2)
(1)
(3)
(1)
(3)
Pfu 1 : ----------MIEDVDYITEESKPDIPLFKK----ENGKFKIEHDRTFRPYLYA----- : 40
Tgo 1 : ----------MIEDTDYITEDSKPDIRIFKK----ENGEFKIDYDRNFRPYLYA----- : 40
PKOD 1 : ----------MIEDTDYITEDSKPDIRIFKK----ENGEFKIKYDRTFRPYLYA----- : 40
DTok 1 : ----------MIEDADYITEDSKPDIRIFKK----ENGEFKIDYDRDFRPYLYA----- : 40
9°N-7 1 : ----------MIEDTDYITENSKPDIRIFKK----ENGEFKIEYDRTFRPYLYA----- : 40
Tli 1 : ----------MIEDTDYITKDSKPDIIRIFKK----ENGEFKIELDPHFQPYLYA----- : 40
Mvo 1 : ----------MQLDENSKD--LCIDMYYKNCGLEKPEINLQKRCPKCNKFYVDTSEP : 45
Pis 1 : ------MELKVWPEDITYAVVGSVPEIRIEGI---LSSGERVVLIDRSKKDFYI----- : 45
Rfu 1 : -----MERVEGWLIDADYETIGKSAIVRLWCK---DDQG-IFVRYDYNKDFRAV----- : 46
Csy 2 : TVQDAVEIPPSLLVSRTYDSQRFAVLLKFYEP---ESQKIVHWTDNTGHKFQST----- : 53
Sac 40 : WIKEARDGKVYFLLQVDVDGKKSRAICKLYDK---EGKKIYIDQDESGHKFELT----- : 91
Soh 41 : WIKEAREGKSYFLIQVDVDGKKSRAICKLYDK---ETKKIYILYDNTGHKFELT----- : 92
Sso 41 : WLEEAQENKIYFLIQVDVDGKKSKAICKLFDK---ETQKIYALYDNTGHKFELV----- : 92
Poc 77 : KPLEARDGVEGELIQTMDGKRFVARAKIYDD---RNGIVYVYEDPTGYMDYELT----- : 128
Xpe 121 : VREPWVESVRGYLIDVEADGSLFAQLLMLYDP---SSGSLVFVWADETGHKPYELT----- : 172
(3)
Pfu 41 : ----------LLRDDSKIEEVKKITGERHGKIVRIVDVEEVKKFFG--------- : 77
Tgo 41 : ----------LLKDDSAIEDVKKITAERHGTTVRVTRAEEVKKKFFG--------- : 77
PKOD 41 : ----------LLKDDSAIEEVKKITAERHGTVVTVKRVEEVQKKFFG--------- : 77
DTok 41 : ----------LLKDDSAIEDIKKITAERHGTTVRVTRAERVKKKFFG--------- : 77
9°N-7 41 : ----------LLKDDSAIEDVKKVTAKRHGTVVKVERAEEVQKKFFG--------- : 77
Tli 41 : ----------LLKDDSAIEEIKAIKGERHGKTVRVLDAVEVEKKFFG--------- : 77
Mvo 46 : KEIYDYLDGLNQKIDLKKLEPEFENNTSLKVQDLITNIKIIEKIVYSDYIINGKDISEVS : 105
Pis 46 : ----------VDCAVCEPARLKTALS-RVAPIDDVQIVERRFFG--------- : 78
Rfu 47 : ----------IGVDEDILKNRATSTER--EVIKLESFEESAQLKTFG--------- : 80
Csy 54 : ----------R-QPPSKLGEL----EGREDVLGTEQVMRHDLIRDK--------- : 84
Sac 92 : ----------D-IDPDKVNKI-TKVVRDPSEDHLELINSVDPYTGK--------- : 125
Soh 93 : ----------D-IDPEKVNKI-PKVVRDPSEDHLETVIKIDPYSGN--------- : 126
Sso 93 : ----------D-LEPDKVGKI-PKIVRDPSEDHIETVSKIDPYTVN--------- : 126
Poc 129 : ----------D-IPPDKLQEL-KEVVRHKGEDHVEVVEKEDLLEVQ--------- : 162
Xpe 173 : ----------D-ARPEDLRRAGVDVSHDESELQYDLVEKFHPIDRK--------- : 207
(2)
Pfu 78 : -----KPITVWKLYLEHPQDKP-TIEEKVREHPAVVDIEYDIPIRKRLIHKCLIPMEG : 131
Tgo 78 : -----RPIEVWKLYFTHPQDKP-RIEDKIKEHPAVVDIYYDIPIRKRLIHKCLVPMEG : 131
PKOD 78 : -----RPVEVWKLYFTHPQDKP-RIEDKIREHPAVIDIYYDIPIRKRLIHKCLVPMEG : 131
DTok 78 : -----RPVEVWKLYFTHPQDKP-RIEDRIREHPAVVDIYYDIPIRKRLIHKCLIPMEG : 131
9°N-7 78 : -----RPIEVWKLYFNHPQDKP-RIEDRIREHPAVVDIYYDIPIRKRLIHKCLIPMEG : 131
Tli 78 : -----REVEVWKLIFEHPQDKP-RMRGKIREHPAVVDIYYDIPIRKRLIHKCLIPMEG : 131
Mvo 106 : DFKNKKERKICKVYVKYENHEK-IIEKYFEEEG---KSYEEIDPILRKCMIQQDIVPSRK : 161
Pis 79 : -----RSKKFLEVIAKIPEDVR-KLEEAAMSIPRVSGVYKKIDPEYMLCCMHMGVVPCSW : 132
Rfu 81 : -----REVEGYIVYAHHPQHKP-KLED---YLSQEGDVRKAIDPIAYLCLIHKDIACMDG : 131
Csy 85 : ------DIPVTKITVRDLLRIGGTNSEKSIENI--MDTWKSDHKYYENLLYFKSLVVGEY : 136
Sac 126 : ------KIRLTKIVVKDLLAVR-RMRSSLPKA------YKAHHKYYNNLYYLNLIPGLI : 172
Soh 127 : ------KIKLTKIVVKDLLAVR-RMRNSVPKA------YKAHHKYFNNLYYLNLIPGLP : 173
Sso 127 : ------KFKLTKIVVRDLLAVR-RLRNDVPKA------YKAHHKYFNNLYYLNLIPGLM : 173
Poc 163 : ------PRKVTKIVVKTDDVVR-VLRDKVPRA------YKANHKYHNNLYYLYLIPGLK : 209
Xpe 208 : ------LVKLTKIVVSDLLAVR-RLREKVSSAG--FSVWEADLKYHDNLIFLRQLIPGLL : 258
(2)
(1)
(2)

Figure 2

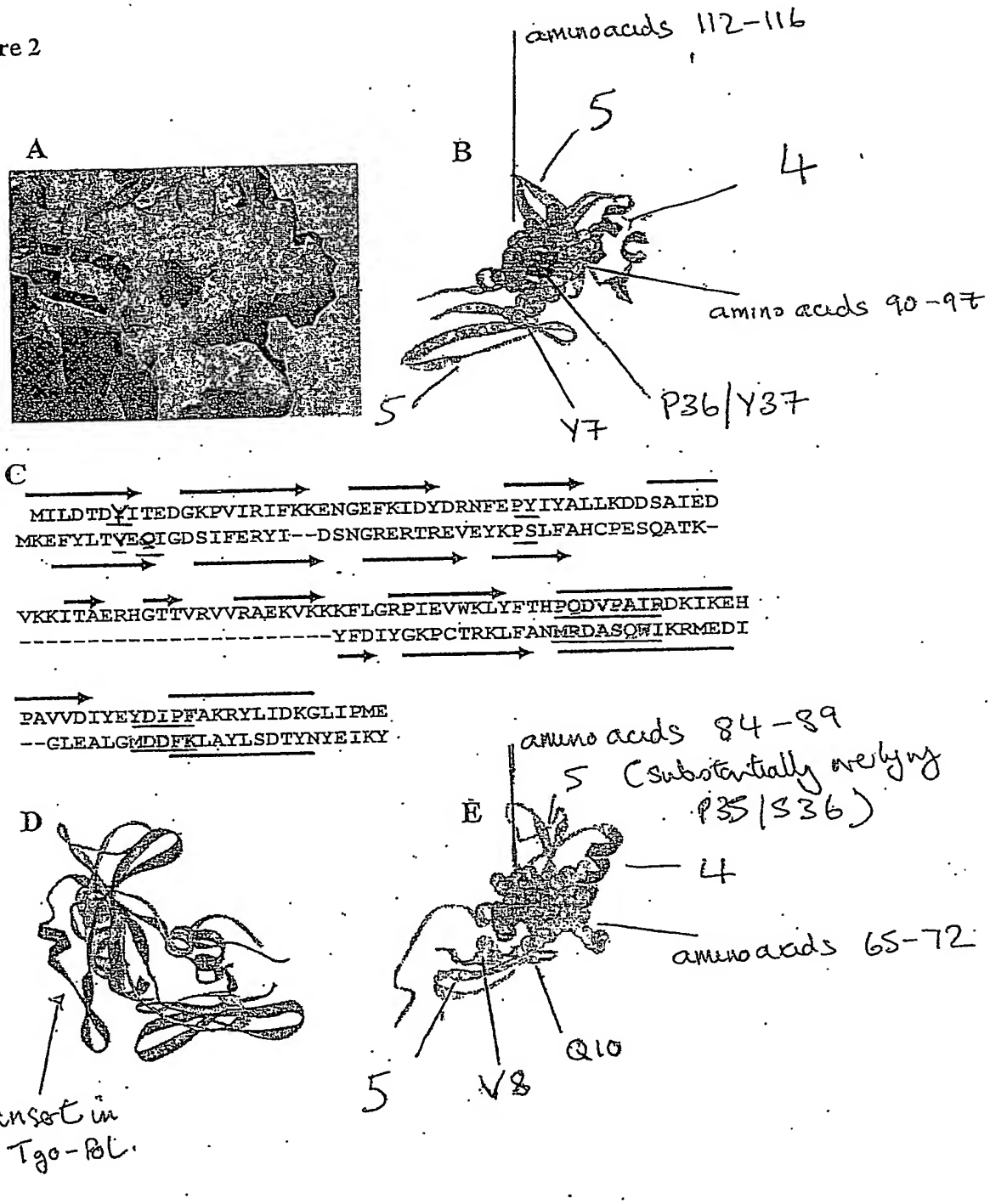

A
B
amino acids 112-116
5
4
amino acids 90-97
5
Y7
P36/Y37
C
MILDTDYITEDGKPVIRIFKKENGEFKIDYDRNFEPYIYALLKDDSAIED
MKEFYLTVEQIGDSIFERYI--DSNGRERTREVEYKPSLFAHCPESQATK-
VKKITAERHGTTVRVVRAEKVKKKFLGRPIEVWKLYFTHPQDVPAIRDKIKEH
----------YFDIYGKPCTRKLFANMRDASQWIKRMEDI
PAVVDIYEYDIPFAKRYLIDKGLIPME
--GLEALGMDDFKLAYLSDTYNYEIKY
D
E
amino acids 84-89
(substantially overlying P35/S36)
5
4
amino acids 65-72
5
V8
Q10
insert in Tgo-PolL

Figure 3

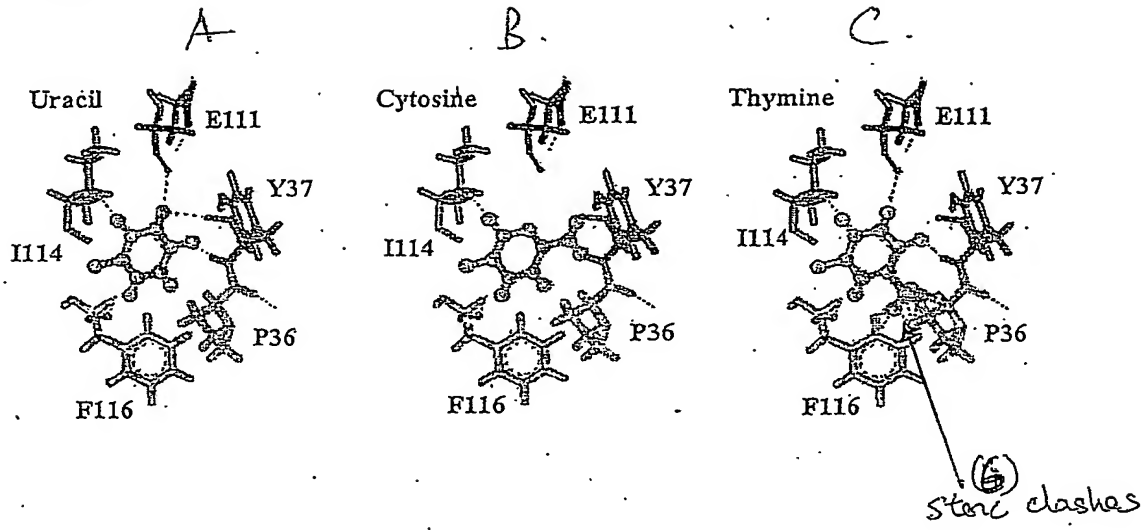
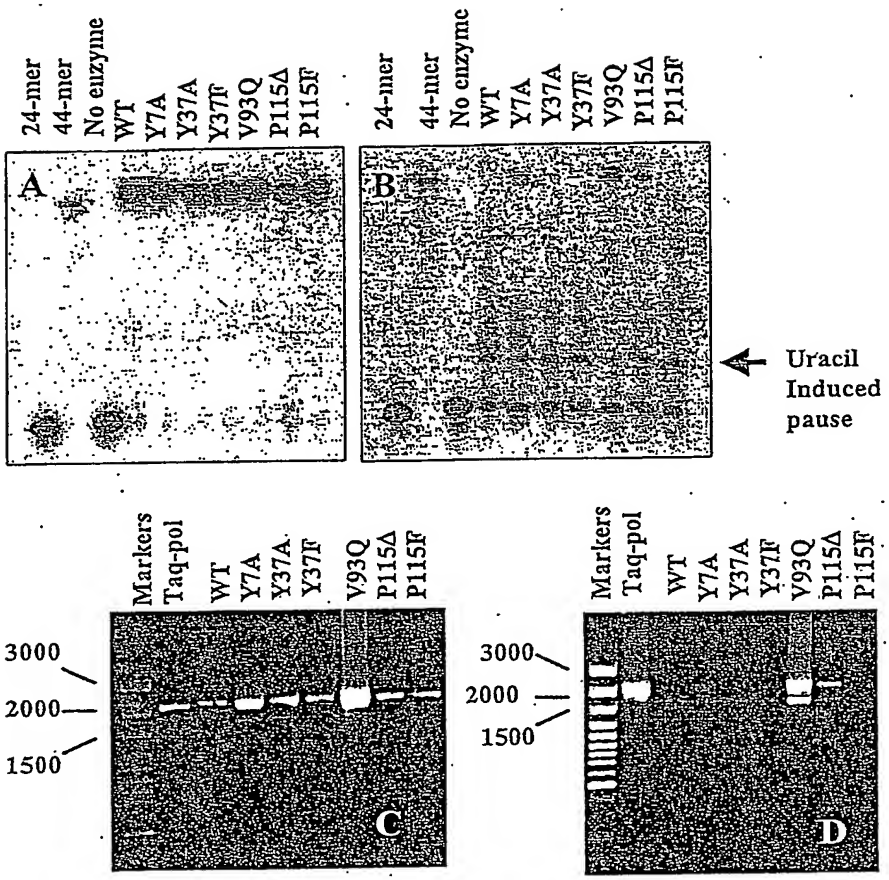

A
B
C
Uracil
E111
Y37
I114
P36
F116
Cytosine
E111
Y37
I114
P36
F116
Thymine
E111
Y37
I114
P36
F116
(6)
steric clashes

Figure 4.


24-mer
44-mer
No enzyme
WT
Y7A
Y37A
Y37F
V93Q
P115Δ
P115F
A
B
Uracil
Induced
pause
Markers
Taq-pol
3000
2000
1500
C
D